

Checklist for appraisal of study relevance (child sex offenses)

First author, year, reference number				
---	--	--	--	--

Relevance	Yes	No	Cannot answer	Not applicable
1. Study population				
a) Is the population from which the participants were selected clearly described and relevant?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Were acceptable procedures applied to recruit participants?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Are the inclusion criteria adequate? ¹	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d) Are the exclusion criteria adequate?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Summary 1 a) – 1 d): Is the study population relevant?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. Test intervention				
a) Is the test intervention one of those previously specified? ²	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Was the test intervention administered/performed in a correct and reproducible manner?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Summary 2 a) – 2 b): Is the test intervention relevant?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Comparison intervention				
a) Is the comparison intervention one of those previously specified? ³	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Is it possible to exclude that the choice of comparison intervention, dose, or method has introduced a systematic error which would favour either intervention?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Summary 3 a) – 3 b): Is the comparison intervention relevant?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Effect measure				
Are relevant effect measures applied in the study? ⁴	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Study duration				
Does the study have an adequate follow-up time? ⁵				

¹Population

Convicted of child sexual offending
Self-reported sexual activity involving prepubertal and early pubertal children
Convicted of child pornography offending
Convicted of other sexual offending

²Test intervention

Pharmacological
Psychological/psychotherapeutic
Combinations of the above

³Comparison intervention

Conventional treatment
No active treatment

⁴Outcomes

Conviction of child sexual offending
Police arrests on suspicion of child sexual offending
Breach of conditions following sentences for sexual offenses
Self-reported child sexual offending
Self-reported sexual impulses which include children
Sexual offending against adults

⁵Study duration

Follow-up at least one year after completion of intervention

Critical Appraisal Form: Randomised Controlled Trials

Summary of *critical appraisal*

Author, year, or SBU identification number:

Overall evaluation of study quality:

High ☐ Moderate ☐ Low ☐

Instructions:

The alternative “unclear” is used when the information was not forthcoming in the text.

The alternative “not applicable” is used when the question is irrelevant.

Some questions have clarifying comments presented as footnotes.

Study quality	Yes	No	Unclear	Not applicable
<i>1. Study population</i>				
a) Does the study state how many individuals were excluded before randomisation?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Does the study adequately account for those who were not randomised, although they qualified for inclusion?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>2. Distribution of measure/intervention/treatment</i>				
a) Was the method of randomisation applied in such a way as to acceptably minimise the risk of manipulation?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Was randomisation carried out in such a way that the distribution was unpredictable and random? ¹	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Did all participants who were randomised begin treatment? ²	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>3. Comparability (similarity) of groups</i>				
a) Were the groups reasonably similar at baseline, with respect to characteristics which can influence the results (e.g. age, sex, severity of illness)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>4. Blinding (masking)³</i>				
Were the following blinded satisfactorily?				
a) Patients	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Study quality	Yes	No	Unclear	Not applicable
b) Those who administered the treatment (operators)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Those who evaluated the results (observers)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>5. Attrition (loss to follow-up) (the number of randomised participants who had not been followed in accordance with the study protocol)⁴</i>				
a) Is it possible to follow the progress of the participants through the study e.g. by means of a flow chart?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Is the level of attrition after randomisation acceptable?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Is the attrition adequately accounted for?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>6. Compliance, adherence, concordance⁵</i>				
a) Does the study state to what extent participants completed the treatment?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Did an acceptable proportion of participants complete the treatment?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>7. Reporting of effectiveness and side effects</i>				
a) Was the primary outcome (measure of effectiveness) defined beforehand and adequately reported?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Were the secondary outcomes (measures of effectiveness) defined beforehand and adequately reported?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Were the conclusions based solely on previously defined outcomes (measures of effectiveness) and analyses of subgroups? ⁶	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d) Have the outcomes of all important measures of effectiveness been adequately presented?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e) Were side effects/complications reported satisfactorily?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>8. Results and precision</i>				
a) Were the results adequately presented? ⁸	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Have the results been calculated using an appropriate method of analysis? ⁹	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Was the minimum clinically relevant effect defined beforehand?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d) Is the selected minimum clinically relevant effect of appropriate magnitude?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Study quality	Yes	No	Unclear	Not applicable
e) Have acceptable methods been applied to measure the outcomes?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f) Was inter-observer agreement evaluated in an acceptable way? ¹⁰	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
g) Are the factors and calculations used to determine the minimum number of participants acceptable (power analysis)? ¹¹	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. Conflicts of interest				
a) Have potential conflicts of interest been disclosed?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Are you convinced that the study results have not been influenced by conflicts of interest?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Overall assessment of study quality				
<input type="checkbox"/> High <input type="checkbox"/> Moderate <input type="checkbox"/> Low				

Comments/footnotes to **critical appraisal form for RCT**

1. The risk that randomisation will be predictable to the observer or the participants can occur, e.g. with block randomisation used in multicentre studies to counteract random, uneven distribution between different centres or countries.
2. This heading determines the risk that the results have been influenced by selective exclusion of participants from the study after randomization, but before treatment start. The number of participants who failed to complete the study should be considered in relation to the size of the study. If the number is evenly distributed between the groups and the reasons presented are acceptable, then the risk that the results have been compromised is minor. If more than 5% of the randomized participants have been lost to follow-up, or if no reason is given for the attrition, or if the reasons given are not acceptable, then the risk is considered to be major.
3. It is preferable that both participants and observers in a study are blinded. For practical reasons it can sometimes be difficult or impossible to conceal from the observer/operator and/or subject which treatment that is given. However, in most cases it is possible to ensure that the observer, the person evaluating the effect of the intervention, is blinded.
The following alternatives are available:
 - Open testing: no party is blinded
 - Single-blind: a) the participants are blinded; b) the operator and/or the observer (the person evaluating the results) is blinded
 - Double-blind: a) the participants and the operator and/or the observer are blinded and the study description affirms that the observations were recorded before the test code, identifying test and control subjects, was broken.

There are numerous examples of studies where blinding has been unsuccessful because of

characteristic effects or side effects of active intervention, such as mouth dryness associated with administration of neuroleptic agents and uterine bleeding associated with oestrogen treatment. In some cases it is possible to administer preparations which counteract the side effects, in order to reduce the risk of compromising the blinding. Other factors which can make blinding difficult are differences between tablets, inhalant compounds etc. with respect to appearance or taste. A pronounced 'placebo-effect' in the control group can indicate successful blinding. In some studies the participants are asked to guess whether they have received active or control treatment.

4. The attrition assessed here refers to subjects who drop out of the study after randomisation. There may, however, be occasions where even considerable attrition is probably coincidental. The examples presented below should therefore be regarded as general guidelines:
 - Small (<10%)
 - Medium (10–19%)
 - Large (20–29%)
 - Very large ($\geq 30\%$). Such a large loss potentially invalidates the results, which can indicate that the study should be excluded.

Attrition varies at different time points in a study and can vary with respect to different outcome measures. Loss to follow-up often increases over time. Therefore the validity of treatment results recorded at the final follow-up event may be doubtful, whereas the results from earlier follow-ups may be valid.

5. Keeping note of participant compliance is especially important in cases where statistical analysis discloses no significant difference in outcomes between the two groups. Poor compliance can reduce both the effects of the intervention and side effects. If the intervention shows a significant effect, records of compliance are less important. The exception is in studies where compliance is poorer in the group that received reference treatment. This can occur in a placebo controlled study if blinding was inadequate, or if a reference treatment has a much higher frequency of side effects.

A guide for acceptable compliance is that more than 80% of the subjects participated in more than 80% of the treatment.

6. It is not unusual for studies with negative results to include explanatory or post hoc analyses, in order to identify certain subgroups in the study sample which have benefited from the intervention. These analyses can have an important function in generating hypotheses, but there is of course a great degree of uncertainty. Study conclusions must therefore never be based on such analyses.
7. Even when the reported outcome is reasonable, defined beforehand and adequately reported, there can be other important outcome measures which have been omitted. Most frequently this applies to the outcome measure for risk assessment, which is also assessed under footnote 8.
8. The usual measurements for dichotomous variables are the relative risk (RR), odds ratio (OR), or absolute risk reduction/risk difference and number needed to treat (NNT). For continuous variables the difference in means, mean difference, is usually used. All such measures should be presented with an appropriate measure of dispersion, preferably with a 95% confidence interval.
9. The results can be analysed according to Intention-to-treat (ITT) and/or per protocol (PP). An ITT analysis means that all subjects who have been randomised are followed up within the frame of

the study, regardless of whether they have been assigned to the treatment group or not. This is often the method of choice. If the results are calculated in other ways there is a risk that the treatment effect will be overestimated. ITT analysis can be complemented with a sensitivity analysis according to the “worst case scenario” in which subjects lost to follow-up from the group showing the best results are included, but assigned the worst possible outcome and those lost to follow-up from the group with the worst outcome are assigned the best possible outcome.

Sometimes it is desirable for a PP analysis to be presented, which means that only those subjects who have followed the entire study protocol are included in the analysis. In the event of attrition in studies using continuous variables or rating scales, occasionally a calculation method is used in which the most recent results are considered to apply even for later time points for which data are unavailable; last observation carried forward (LOCF).

10. In registering the outcomes in a treatment study, interobserver variation can be a weakness (source of error), for example in studies where several observers are to evaluate radiographs or cytology samples. In such cases, interobserver agreement among most or all of the observers should be reported. This can be expressed in the form of a Kappa coefficient, or Intraclass correlation coefficient (ICC), depending on which scale is used.
11. Power calculations are used to calculate the statistical strength of a study, i.e. to calculate beforehand how many subjects should be included in order to demonstrate a treatment effect with reasonable probability. It is important that the authors describe how they have arrived at the selected sample size and that the calculations have been done prior to study start. Otherwise it is impossible to rule out the likelihood that the authors have successively added subjects to the study until statistical significance was achieved.

Critical appraisal form: Cohort studies with control groups

Summary of *appraisal*

Author, year, or SBU identification number:

Overall evaluation of study quality:

☐ High ☐ Moderate ☐ Low

To be used for:

Evaluating the effect and safety of interventions.

Evaluating the importance of risk factors/risk markers in predicting disease.

The terminology can vary, but in all cases an intervention group (synonyms: exposed group, cases or risk factor group) is compared with a control group (synonyms: unexposed group, comparison or reference group).

1. Comparability/similarity	Yes	No	Unclear	Not applicable
<i>1.1 The groups being compared</i>				
a) Have the compared groups been adequately selected? ¹	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Is the control group relevant?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Is it likely that the intervention and control groups were selected and diagnosed in a similar manner? ²	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>1.2 Group comparability (similarity and confounders)</i>				
a) Have the authors identified all important confounding factors (see below)? ³	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Have the authors taken these factors into account in their analyses? ³	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Were any differences in baseline characteristics negligible (see confounding factors listed below)? ³	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d) Is the risk of selection or indication bias small? ⁴	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>1.3 Intervention</i>				
a) Is the intervention clearly defined with respect to content and quality?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

b) Is the intervention in the comparison group clearly defined with respect to content and quality?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<p><i>Confounding factors</i></p> <ul style="list-style-type: none"> • age • previous convictions for sexual offences • non-contact sexual offences • previous violence against a person • other criminality • relationship to victim (known/unknown) • the sex of the victim • stable adult relationships • for historical controls – time aspects 				
2. Compliance, attrition				
<i>2.1 Compliance, adherence</i>				
a) Does the report disclose the proportion of participants who completed the treatment?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Was the proportion completing treatment acceptable?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
<i>2.2 Attrition (loss to follow-up) (number of participants not followed up in accordance with study protocol)</i>				
a) Is the magnitude of attrition (loss to follow-up) presented? ⁵	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Are the reasons for loss to follow-up presented? ⁵	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Is this level of attrition acceptable? ⁵	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. Blinding				
Were the observers (those responsible for evaluating the outcomes) unaware of whether the subject belonged to the intervention or the control group? ⁶	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Statistical power				
a) Is there a clear description of the factors and calculations on which the minimum sample size was determined? ⁷	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Is the statistical power high enough? ⁷	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. Effect measure and statistical analysis				
a) Are individuals with a primary effect measure adequately identified?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Is there only minor risk of recording or measurement bias?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c) Has the statistical analysis of reliability been adequately managed? ⁸	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

d) Have the authors adequately corrected imbalances between the groups with respect to confounders? ⁹	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e) Have treatment drop-outs been taken into account?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. Side effects				
Were side effects/complications measured satisfactorily?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. Conflicts of interest				
a) Does the report include a list of potential conflicts of interest?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b) Are you convinced that the study results have not been influenced by conflicts of interest?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Overall evaluation of study quality:				
<input type="checkbox"/> High <input type="checkbox"/> Moderate <input type="checkbox"/> Low				

Comments on the *critical appraisal form for cohort studies with control groups*

In studies designed as cohort studies with control groups, at least two groups are followed longitudinally into the future, in order to observe what happens to them. This can apply to both non-randomised control studies and other observational studies in which either treatment measures or risk factors are studied.

Synonymous terms are:

Intervention group = exposed group = risk factor group = cases

Control group = unexposed group = comparison group = reference group

1. Is the comparison group clearly defined? Was the intervention compared with another intervention or with no intervention at all? Was the comparison group sampled from the general population or from a limited, selected population? If the comparison group is a historical control particular caution is warranted in appraisal of the study.
2. An important question is whether the same methodology was used to assign subjects to the intervention and control groups, respectively.
3. Confounders are background variables which influence the outcome. They can be unevenly distributed between the groups and thus compromise the "true" result. Among important confounders are age, sex, underlying history of disease, concurrence of several diseases, or risk factors and not the least, socioeconomic status. Socioeconomic status is probably the greatest risk factor for ill health and premature death.

Information that could disclose pronounced differences between groups is usually presented in an introductory table of baseline characteristics.

4. Selection bias occurs when there are one or several intrinsic differences between the groups which may explain the results. The risk is especially high with respect to preventive measures or measures to alleviate symptoms, which well-informed patient groups may request. The risk of selection bias is also high if the intervention is particularly appropriate for application in high- or low-risk patients.
5. High attrition generally increases the risk that results are compromised by systematic errors. Cases arise, however, when even a high level of attrition is probably random/ coincidental. As a general guideline in drug studies, the risk is minor if attrition is less than 10%, medium if attrition is between 10 and 19% and high if attrition is between 20 and 29%. If the attrition in drug studies is 30% or more then the losses may potentially invalidate the study and it may be excluded. Attrition can vary between different time points and with respect to different outcome measures. In studies with long term follow-up, a somewhat higher level of attrition may be acceptable.
6. If the observers are aware of which treatment the subjects have received this can increase the risk of systematic errors in registration.
7. Small studies in which the researchers did not calculate beforehand the minimum sample size required to achieve a statistically significant result for the primary outcome often have major shortcomings with respect to quality. It is important to assess the study's statistical power for each individual outcome measure. An example is the reporting of side effects. Studies are usually planned to highlight the positive effects and may not have taken into account the minimum number of participants required to achieve statistically confirmed negative effects.
8. Assess whether the confidence intervals or other relevant measures are adequately presented or if there is an explanation to why such information has not been presented. This can apply for example to total examinations of large sets of data.
9. Methods that can be applied in this context are matching/restriction, stratified analysis, multivariate model analysis (e.g. regression analysis) or propensity score-methods